DOCUMENT RESUME

ED 021 843 TE 000 558

THE STATE OF WRITTEN COMPOSITION IN RICHFIELD SENIOR HIGH SCHOOL: REPORT OF AN EVALUATIVE STUDY CONDUCTED BY THE LANGUAGE ARTS DEPARTMENT.

Richfield Public Schools, Minn.
Pub Date 68
Note- 50p.
EDRS Price MF-$0.25 HC-$2.08
Descriptors- COMPOSITION (LITERARY), *COMPOSITION SKILLS (LITERARY), *CURRICULUM EVALUATION, *ENGLISH CURRICULUM, ENGLISH INSTRUCTION, GRADING, HIGH SCHOOL STUDENTS, LANGUAGE, LANGUAGE ARTS, LANGUAGE SKILLS, LANGUAGE TESTS, MEASUREMENT TECHNIQUES, *SECONDARY EDUCATION, SECONDARY SCHOOL STUDENTS, *TEACHER AIDES, TESTING, WRITING SKILLS

The Richfield, Minnesota, High School Language Arts Department studied the composition skills of 24 classes in grades 10, 11, and 12 to assess (1) improvement in writing skills from each grade level to the next as an indication of the effectiveness of a composition curriculum which emphasized expository writing, a d (2) the effects of lay readers on the composition program. The Sequential Tests of Educational Progress Writing Test, Form 2A, and an impromptu expository composition were used as measures of student achievement. The lay-reader program was evaluated through questionnaires and a comparison of the achievements of classes with and without lay readers. Results indicated that both average-above average students and average-below average students improved from grade to grade, that the mean and median scores of these students were well above national norms, and that lay readers were not a liability to the program but received approval from both students and teachers. (JS)

THE STATE OF WRITTEN COMPOSITION IN RICHFIELD SENIOR HIGH SCHOOL: REPORT OF AN EVALUATIVE STUDY

CONDUCTED BY

THE LANGUAGE ARTS DEPARTMENT

TE 000 558
ED021843

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

THE STATE OF WRITTEN COMPOSITION IN RICHFIELD SENIOR HIGH SCHOOL: REPORT OF AN EVALUATIVE STUDY CONDUCTED BY THE LANGUAGE ARTS DEPARTMENT

Richfield Public Schools

Richfield, Minnesota

1968

SUPERINTENDENT OF SCHOOLS

Carlton W. Lytle

BOARD OF EDUCATION

Robert B. Gunderson

Richard A. Carlson

James C. Swanson

Lonnie E. Dye

Richard F. Everson

Robert W. Wagner

DIRECTOR OF SECONDARY EDUCATION

Harold A. Rasmussen

SENIOR HIGH SCHOOL PRINCIPALS

Clair Bartosh

Alfred Krinke

Gene Olive

SENIOR HIGH SCHOOL DEPARTMENT CHAIRMEN

Lucille Duggan, 1966-67
Joyce Schwartz, 1967-68

EVALUATION COMMITTEE CHAIRMAN

Harold S. Webster

DISTRICT CHAIRMAN

Lillian B. Ford

FOREWORD

It is with a deep sense of satisfaction that we express commendation to the language arts department and their district chairman, Mrs. Lillian Ford, for their extensive efforts in this timely evaluative effort.

The REPORT OF AN EVALUATIVE STUDY ON THE STATE OF WRITTEN COMPOSITION represents a step forward in the educational development process in our school district. It may well serve as a model and inspiration for future efforts. Well designed and implemented educational research continues to be a neglected facet of the development process. The need is very great if we are to make prudent and effective decisions for change. It is indeed very satisfying when local staff members not specifically trained in sophisticated evaluative procedures undertake a task of this magnitude. It is particularly significant coming on the heels of a long-range comprehensive language arts curriculum development program.

Carlton W. Lytle,
Superintendent of Schools

Harold A. Rasmussen,
Director of Secondary Education

TABLE OF CONTENTS

Introduction i

Purposes of the Evaluative Study 1

Sources of Data Used in the Evaluation 1

The Design of the Testing Program 1
The Nature of the Measuring Instruments Employed 1
Selection of the Test Sample 3

Tabulation and Interpretation of Scores Achieved by Students at Various Grade-Levels 5

Tabulation and Interpretation of Scores Earned by Students in Average-Above Average and Average-Below Average Groups 10

Comparison of Local Scores with National Norms 12

Analysis of Teachers' Ratings of Test Compositions 16

Assessment of the Effects of the Use of Lay Readers to Assist Teachers in the Evaluation of Student Compositions . . . 19
Tabulation and Interpretation of Relevant Test Data 19
Interpretation of Responses to Questionnaires 25
Participating Teachers' Assessment of the Program 33

Summary of Conclusions Based on This Study 36

Recommendations Stemming from the Findings of This Study 37

Appendix
Instructions for Impromptu Test Composition 40
Student Questionnaire 41
Counselors' and Administrators' Questionnaire 44

INTRODUCTION

The study here reported was initiated and carried out within the language arts department in the spring of 1967. Its design was adapted from a plan conceived by Paul Diedrich of Educational Testing Service, to whom we acknowledge our indebtedness.

This report represents a total department effort; it summarizes a project to which every English teacher contributed in some way. All twenty-six participated in the process leading to the selection of a suitable topic for the impromptu compositions and in their subsequent rating. For the nineteen with classes in the sample population, additional meetings concerning procedures were involved.

Those utilizing lay reader service willingly assumed further responsibilities. Serving on the committee to devise questionnaires that would elicit from students, counselors and administrators significant reactions to the lay reader program were Lucille Duggan, Joan Larson, Carol Lohse, Delores Ostgaard, Joyce Schwartz, and Peter Sterling. Harold Webster, as chairman for the study, devoted many hours to his task; he did an outstanding job in working with the district chairman to plan and carry out the necessary procedures. This involved preparing instruction sheets, designing forms, handling mechanical details required for smooth operation of the plan and keeping records that will make it simpler to carry out any evaluative study the department may undertake in the future.

Lillian B. Ford

Lillian B. Ford
District Chairman for Language Arts

PURPOSES OF THE EVALUATIVE STUDY

1. To assess students' progress in writing ability during the senior high school years as an indication of the effectiveness of our structured curriculum in written composition

2. To assess the effects of the use of lay readers to assist teachers in the evaluation of student compositions

SOURCES OF DATA USED IN THE EVALUATION

1. Data related to the first purpose were compiled from students' scores on tests designed to measure achievement in written composition

2. Data related to the second purpose were drawn from these same test results and also from two other sources:

   Questionnaires completed by students in classes involved in the lay reader program and by the counselors and principals concerning reactions to this program

   Evaluative statements made by teachers involved in the program.

THE DESIGN OF THE TESTING PROGRAM

THE NATURE OF THE MEASURING INSTRUMENTS EMPLOYED

Two tests were administered. One was a locally devised performance test, an impromptu expository composition requiring application of the skills emphasized in our written composition curriculum. For this first attempt to carry out an evaluative study, department members adopted the basic pattern of a tested design for administering and rating these test themes. The plan was devised by Paul Diederich and reported in the April, 1966 ENGLISH JOURNAL.

All students wrote on the same topic, which had been selected from among more than seventy possibilities suggested by the teachers but which remained unknown to both teachers and students until the time of writing. In the interests of uniformity, instructions for writing and regulations concerning testing conditions were supplied in printed form.

All members of the department participated in the evaluation of these test compositions, each of which was rated twice, each time by a different teacher. Our concern was with group rather than individual achievement and with students' rather than teachers' performance. We also wished to eliminate any possibility that a student's identity might influence the rating of his paper or that one teacher's opinion might influence another's. Complete anonymity of both writers and raters was therefore maintained.

The test compositions were divided into sets of approximately 60 papers, each set representing a mixture of various categories of students at various grade levels. No information concerning the writer appeared on any composition; each paper bore only a student identification number. Thus no rater could know which student had written any composition; further, he could not know the student's ability level or even whether he was a sophomore, junior or senior.

Department members were divided into two rater groups roughly equated in terms of teaching experience, sex, and grade levels taught. Each was assigned an identification number unknown to the other raters. The first teacher to rate a set of compositions made no marks on the papers themselves; he recorded his rating for each on a separate card on which he also entered the student's identification number and his own. The second rater received only the set of papers unaccompanied by the cards containing the first teacher's ratings. Thus he was unaware of what rating any paper had received and even of which teacher had previously rated the set.

Raters were instructed to judge the compositions holistically--that is, on the basis of their total impact--and to rate them high, middle, or low in terms of their success in meeting the requirements of the assignment and the relative effectiveness of the resulting composition. They were required to observe the arbitrary percentages specified in the Diederich plan, assigning a high (H) rating to 25% of the papers in the set, a medium (M) rating to the middle 50%, and a low (L) rating to 25%. In order to maintain these fixed proportions, they first divided the papers into three groups according to quality and then, if necessary, adjusted the size of the groups by moving some borderline cases up or down one category. In order to facilitate later computations, numerical values were arbitrarily assigned to these letter ratings: H = 6; M = 4; L = 2.

The final rating for each paper was then represented by one of the six possible combinations of letters or the sum of the numerical values of the two letters: HH (12); HM (10); MM (8); LM (6); LL (4); HL (8). This was a departure from the Diederich plan, which involves a reconciliation of differing ratings through a third rating by some particularly competent rater or through consultation with the writer's teacher. Use of such a composite rating not only avoids the introduction of personalities into a situation designed to eliminate this factor but also, according to studies reported by Godshalk, Swineford and Coffman in the 1966 College Entrance Examination Board Monograph THE MEASUREMENT OF WRITING ABILITY, provides a more accurate measure of achievement. However, the HL rating seemed unacceptable. By numerical value, papers so rated would be equated with those receiving two M ratings; yet it seemed unlikely that the papers in these two groups would actually be of the same general quality. When test results were analyzed, the scores for the small number of papers receiving these widely disparate ratings (25 of 659) were therefore ignored as undeserving of confidence.

The other measuring instrument employed was the STEP (Sequential Tests of Educational Progress) Writing Test, Form 2A. This is a standardized

objective test of written composition skills designed for use at the senior high school level. It is sufficiently respected by the National Council of Teachers of English to be one of the measures used in its annual competition for outstanding achievement awards. STEP scores provide a basis for comparing the achievement of students in various categories at the various grade levels within our own school. STEP norms tables also provide a basis for comparing the achievement of Richfield students with that of students in schools throughout the nation.

The use of these two testing instruments is in line with the statement by Edward S. Noyes of Educational Testing Service in his foreword to THE MEASUREMENT OF WRITING ABILITY that "a combination of objective items (which measure accurately some skills involved in writing) with an essay (which measures directly, if somewhat less accurately, the writing itself) [is] more valid than either type of item alone."

SELECTION OF THE TEST SAMPLE

Twenty-four class sections, eight at each grade level, were included in the test sample. As nearly as was possible when working with existing rather than specially created classes, the population sample included equal representation of average-above average (I) and average-below average (II) groups and of lay reader (LR) and non-lay reader (NLR) groups. Further, the attempt was made to include an equal representation of average-above average classes having the services of a lay reader (ILR), average-above average classes not having such services (INLR), average-below average classes with such service (IILR) and average-below average classes without it (IINLR). Nongraded classes were excluded from the testing because these slow-learning students had not participated in the same learning experiences as those in classes following the regular curriculum. As will be explained below, this omission was compensated for when comparing Richfield students' achievement on the STEP Test with the national norms.

The number of students in each category at each grade level who actually participated in the testing by each measure appears in Tables 1 and 2 below:

TABLE 1
Sample Population for Written Compositions
(after removing 25 rated HL)

	Total	ILR	INLR	IILR	IINLR	I	II	LR	NLR
Grade 10	216	56	61	52	47	117	99	108	108
Grade 11	210	57	53	49	51	110	100	106	104
Grade 12	208	53	50	27	78	103	105	80	128
Grades 10-12	634	166	164	128	176	330	304	294	340

TABLE 2
Sample Population for STEP Writing Test

	Total	LR	NLR	
Grade 10	227	114	113	(No separate tallies distinguishing I's from II's were made for this measure. Results of this test were originally intended for use only in evaluating the lay reader program.)
Grade 11	230	114	116	
Grade 12	227	83	144	
Grades 10-12	684	311	373	

TABLE 3
Proportionate Representation of Total Populations
Constituted by Sample Groups Tested

	STEP Writing Test		Test Composition
Test Sample	Percentage of Total Grade-Level Population	Percentage of Total Grade-Level Population Exclusive of Nongraded Classes[1]	Percentage of Total Grade-Level Population Exclusive of Nongraded Classes
Grade 10	25.1	26.2	24.29
Grade 11	28.2	29.4	26.8
Grade 12	27.6	28.9	26.5

[1]Representation of nongraded students is irrelevant to a comparison of the achievement at various grade levels by Richfield students following the established curriculum. It is, however, useful for comparing the scores achieved by Richfield students with the national norms.

THE LARGE SIZE OF THE SAMPLE GROUPS INCREASES THE DEGREE OF CONFIDENCE THAT MAY BE PLACED IN THE TEST SCORES AS REFLECTING ACCURATELY THE ACHIEVEMENT OF THE ENTIRE GROUP REPRESENTED BY THE SAMPLE. IT ALSO SERVES TO COMPENSATE FOR THE UNAVAILABILITY OF PERFECTLY MATCHED GROUPS AND, IN CONNECTION WITH THE LOCAL TESTING INSTRUMENT, FOR THE FACT THAT THE ASSIGNED TOPIC MIGHT BE MORE APPEALING FOR SOME STUDENTS THAN FOR OTHERS AND MIGHT THEREFORE INFLUENCE THE QUALITY OF INDIVIDUAL COMPOSITIONS.

Students at each grade level in the sample group had participated for three years in the structured composition program instituted in the secondary schools in the fall of 1964, tenth graders in grades 8-10, eleventh graders in grades 9-11, and twelfth graders in grades 10-12.

TABLE 4
The Number of Teachers With Classes in the Test Sample

	Experienced in Richfield	Experienced but New in Richfield	Inexperienced	Total
Grade 10	6	0	1	7
Grade 11	5	1	1	7
Grade 12	4	1	0	5
Grades 10-12	15	2	2	19

Students tested thus represent those taught, not by a selected minority of teachers, but by 79% of all teachers in the department.

THIS HIGH DEGREE OF TEACHER INVOLVEMENT, TOGETHER WITH THE FACT THAT THE SAMPLES FOR ALL GRADE LEVELS REPRESENTED THREE YEARS' PARTICIPATION IN THE SORT OF LEARNING EXPERIENCES CHARACTERIZING OUR WRITTEN COMPOSITION CURRICULUM, ALSO INCREASES THE LEVEL OF CONFIDENCE TO BE PLACED IN THE TEST RESULTS. TEST SCORES MAY BE EXPECTED TO REFLECT ACHIEVEMENT AS INFLUENCED BY THE NATURE OF THE CURRICULUM RATHER THAN BY DIFFERING TERMS OF STUDENTS' EXPERIENCE WITH IT OR DIFFERING LEVELS OF TEACHER PERFORMANCE.

TABULATION AND INTERPRETATION OF SCORES ACHIEVED BY STUDENTS AT VARIOUS GRADE-LEVELS

Testing of a cross-section of the school population during a single week of a single year cannot tell us about the progress of individual students from year to year, but it can tell us whether students in the various grades achieve at like or different levels. It is this group achievement rather than individual achievement that concerns us when we attempt to discover whether students in general improve their writing skills during their participation in our structured curriculum for written composition. There are a number of ways of looking at the test results in order to determine whether there is evidence of such improvement.

We may examine the range of scores and the percentages of students scoring at various points or within certain intervals of this range at each grade level. This information is recorded in Tables 5 and 6.

TABLE 5
Range and Distribution of Scores on STEP Writing Test

Grade	Sample Size	Highest Score	Upper Quartile[1]	Median[2]	Lower Quartile[3]	Lowest Score
10	227	331	309	297	286	247
11	230	338	311	299	289	247
12	227	342	317	306	295	262

[1]25% of the scores are at or above this point.

[2]50% of the scores are at or above this midpoint.

[3]25% of the scores are at or below this point.

TABLE 6
Range and Distribution of Ratings on Test Composition

Grade	Sample Size	Percentage of Sample Earning Each Combined Rating				
		HH	HM	MM	LM	LL
10	216	10.65	15.74	32.87	20.37	20.37
11	210	11.43	22.38	32.85	17.62	15.71
12	208	20.67	24.52	29.33	13.94	11.54

It will be noted that midpoint and quartile scores on the STEP Writing Test (Table 5) are higher at each succeeding grade level, despite the fact that the lowest individual score at Grade 11 is the same as that for Grade 10.

Table 6 shows the percentage of twelfth graders earning HH ratings as nearly double that for tenth graders, while the percentage of twelfth graders earning LL ratings is just over half that for tenth graders. For LL ratings, the eleventh graders' percentage was about midway between those for tenth and twelfth graders. While the percentage of eleventh graders earning HH ratings is only slightly higher than that for tenth graders, there is a marked increase in the percentage earning HM ratings. The percentage earning MM decreases slightly from grade to grade.

In interpreting the significance of these trends and others to be noted below, it is important to remember a fact pointed out by the originator of the Diederich plan: percentages of students earning each possible rating would be the same at all grade levels if no improvement occurred from year to year. It is therefore encouraging that the combined percentages of those earning HH and HM ratings increases from grade to grade, while a corresponding decrease occurs in the combined

percentages earning LM and LL (Table 8). Perhaps particularly gratifying are the data concerning extremes of achievement: the percentage of twelfth graders earning HH ratings is almost identical with the percentage of tenth graders earning LL, and the percentage of twelfth graders earning LL corresponds rather closely to that for tenth graders earning HH.

THUS FAR, THEN, OUR EXAMINATION OF THE DATA FROM BOTH THE STANDARDIZED OBJECTIVE MEASURE AND THE LOCAL TEST OF ACTUAL WRITING PERFORMANCE INDICATES THAT GENERAL IMPROVEMENT IN WRITING SKILLS DOES INDEED OCCUR FROM EACH GRADE LEVEL TO THE NEXT.

Another way of looking at the test results is to examine the mean, or average, scores earned by students at the three grade levels and the percentages of students in each grade scoring above and below the mean for the entire school sample. This information appears in Tables 7 and 8.

TABLE 7
Grade-Level Means and Comparative Achievement in Relation to the School Mean for the STEP Writing Test[1]

Sample Group	Group Mean	Percentage Scoring Above School Mean (301)	Percentage Scoring Below School Mean (301)
Grade 10	297	37.0	62.9
Grade 11	299	44.8	55.2
Grade 12	306	49.3	50.7
Grades 10-12	301	45.5	54.5

[1]No students scored exactly at the computed mean, or average, which for this measure is a hypothetical score.

TABLE 8
Grade-Level Means and Comparative Achievement in Relation to the School Mean for the Test Composition[1]

Sample Group	Group Mean[2]	Percentage Scoring Above School Mean (800)[2]	Percentage Scoring Below School Mean (800)[2]
Grade 10	751.8	26.39	40.74
Grade 11	792.3	33.81	33.33
Grade 12	857.6	45.19	25.48
Grades 10-12	800.0	35.02	33.28

[1]See MM column of Table 6 for percentages of students scoring at the mean on this measure.

[2]To make possible finer distinctions when computing averages, the values originally assigned to the various combined ratings (HH = 12; HM = 10; MM = 8; LM = 6; LL = 4) were first converted to hundreds: HH = 1200; HM = 1000; MM = 800; LM = 600; LL = 400. There is ample and respectable precedent both for the assignment of arbitrary numerical values to ratings and for such a change in scale. The College Entrance Examination Board converts an original score of 5 to 500 when reporting students' achievement. Paul Diederich, Director of Research in English for Educational Testing Service, also recommends this procedure. In his previously mentioned ENGLISH JOURNAL article (pp. 440, 441), he emphasizes that--far from being an attempt to lie with statistics--such a change of scale is actually a better way of telling the truth, especially when reporting large-group averages, which are notoriously sluggish. "Several students would have to move up a notch [interval of 100] to make a difference of even five points" in a group's average on a scale of 100-900; thus, the larger numbers will "convey to the public a more accurate impression of the growth such differences represent" than would averages computed on a scale of 1-9. Five points on the larger scale represents substantial growth; yet when the same amount of growth is expressed in terms of the smaller scale, this decimal figure (.05) is likely to mislead some people into considering the gain so small as to be insignificant. Although Mr. Diederich's conclusions relate to a study employing a 100-900 scale, they also hold for our study, where the intervals are larger (200 instead of 100) and the number of attainable scores (5 instead of 9) is smaller, but the resulting spread is the same (800 points between the extremes).

On the standardized test, both the group mean and the percentage scoring above the school mean are higher at each succeeding grade level, while the percentage scoring below the school mean decreases from grade level to grade level. The percentage of scores above the school average is larger by 7.8% at Grade 11 than at Grade 10; this is almost exactly balanced by a 7.7% decrease in the percentage of scores below the school average. The percentage of students in Grade 12 who scored above the school mean is 4.5% greater than that for students in Grade 11; the decrease in the percentage scoring below it is an identical 4.5%. The actual point increase in the class mean is, however, greater at Grade 12 than at Grade 11, and the highest and lowest individual scores are less widely separated at Grade 12 than at either preceding grade (Table 5).

On the local measure of performance, the difference between Grade 11 and Grade 12 percentages scoring above the school mean (11.38%) is greater than that between Grade 10 and 11 (7.42%). The decrease in the percentages scoring below the school mean is only slightly smaller at Grade 12 (7.33%) than at Grade 11 (7.41%). The decrease in the percentage of those scoring at the mean is only .02% between Grades 10 and 11, but 3.52% between Grades 11 and 12. The difference in class means between Grades 11 and 12 (65.3 points) is more than half again as large as that between Grades 10 and 11 (40.5 points).

Scores on the test composition were analyzed in still another way as a further check on the trends revealed in the previous tables. The information in Table 9 serves to confirm the previously noted trends.

TABLE 9
Percentages of the Total Number of Scores in Each Rating Category Earned by Students in Each Sample Group

Sample Group	Number in Group	Percentage of Total Number of Earned Ratings Earned by Students in Each Sample				
		90 HH's	132 HM's	201 MM's	110 LM's	101 LL's
Grade 10	216	25.55	25.76	35.32	40.00	43.56
Grade 11	210	26.67	35.60	34.33	33.64	32.67
Grade 12	208	47.77	38.64	30.35	26.36	23.76
Grades 10-12	634	99.99%	100.00%	100.00%	100.00%	99.99%

On the STEP Test the lowest individual scores at Grade 12, though higher than the lowest score for Grade 10, are well below the Grade 10 class average, while the highest Grade 12 scores are well above both the Grade 12 average and the highest score at Grade 10. Though a smaller proportion of students at Grade 12 than at Grade 10 received LL ratings on the test composition, some twelfth graders still achieved this lowest possible score (400).

No coefficient of correlation was computed to indicate the exact relationship of the two test instruments as measures of writing competence. This would have required the pairing of scores earned by individual students, a laborious process deemed unnecessary because we were less concerned with the exact degree of covariance of the two measures than with the growth trends they revealed. If a similar study is carried out in the future, it might be well to arrange in advance for the clerical operations needed to assemble the data on which such a statistical computation could be based. It would be interesting to know how consistently the scores on our local performance test and on the standardized objective STEP Writing Test tend to correspond with each other.

THE TRENDS REFLECTED IN BOTH SETS OF TEST DATA USED IN THIS STUDY SEEM TO LEND SUPPORT TO TWO INFERENCES:

1. IMPROVEMENT DOES INDEED OCCUR FROM GRADE LEVEL TO GRADE LEVEL IN SUFFICIENT DEGREE AND IN THE PERFORMANCE OF A SUFFICIENT NUMBER OF STUDENTS TO SUBSTANTIALLY AFFECT THE CLASS AVERAGE, THE SLUGGISH NATURE OF WHICH HAS BEEN POINTED OUT EARLIER.

2. INDIVIDUAL DIFFERENCES IN ACHIEVEMENT WITHIN A GROUP TEND TO INCREASE RATHER THAN TO DISAPPEAR AS STUDENTS PROGRESS THROUGH OUR THREE-YEAR CURRICULUM--A CIRCUMSTANCE THAT IS USUALLY TAKEN TO MEAN THAT EFFECTIVE TEACHING AND LEARNING HAVE TAKEN PLACE.

These inferences have been based solely on analysis of the achievement of the entire test sample at each grade level. It is also possible to analyze the achievement for subsections of the total grade-level samples to see whether the trends are the same for the average-above average classes (I) and the average-below average classes (II) that make up the total sample at each grade level.

As pointed out earlier, no such tabulation was prepared for the STEP Writing Test scores, since this instrument was originally intended for use only in evaluating the effectiveness of the lay reader program. Relevant information concerning achievement on the test composition appears in Tables 10 and 11.

TABULATION AND INTERPRETATION OF SCORES EARNED BY STUDENTS IN AVERAGE-ABOVE AVERAGE AND AVERAGE-BELOW AVERAGE GROUPS

TABLE 10
Comparative Percentages of I's and II's at Each Grade-Level Earning Each Possible Combined Rating[1]

Sample Group	Sample Size	Percentage of Sample Group Earning Each Rating				
		HH	HM	MM	LM	LL
10 I	117	17.09	23.08	41.03	14.53	4.27
10 II	99	3.03	7.07	23.23	27.27	39.39
11 I	110	21.82	33.64	35.99	7.73	1.82
11 II	100	0.00	10.00	30.00	29.00	31.00
12 I	103	29.13	31.07	26.21	8.74	4.85
12 II	105	12.38	18.10	32.37	19.05	18.10

[1]Corresponding percentages for total grade-level samples appear in Table 6.

It will be noted that the percentage of I's earning HH ratings increases substantially from each grade to the next. For II's the percentage of HH's is about four times as large at Grade 12 as at Grade 10, but there is a drop to zero at Grade 11 from 3.03% at Grade 10. The percentage of II's earning LL ratings at Grade 12 is less than half that for II's at Grade 10, but for I's it is slightly larger at Grade 12 than at Grade 10 and considerably larger than at Grade 11. This circumstance seems to be mitigated somewhat when LM ratings are considered along with LL ratings as representing below average achievement. (See Table 11.)

TABLE 11
Comparative Achievement of I's and II's at Various Grade-Levels in Relation to the Mean for the Entire School Sample

Grade Level	Percentage Scoring Above the School Mean (800)		Percentage Scoring At the School Mean (800)		Percentage Scoring Below the School Mean (800)	
	I's	II's	I's	II's	I's	II's
10	40.17	10.10	41.03	23.23	18.80	66.66
11	55.46	10.00	35.99	30.00	9.55	60.00
12	60.20	30.48	26.21	32.37	13.59	37.15

While the percentage scoring above the school mean at each grade-level is, as would be expected, greater for I's than II's, the percentage earning these above average ratings is greater for both I's and II's at Grade 12 than at Grade 10. It is, in fact, proportionately larger for average-below average classes than for average-above average classes (over 3 times as large for II's and about half again as large for I's). There is also a decrease from Grade 10 to Grade 12 in the percentage of both I's and II's scoring below the school average; for I's this percentage at Grade 12 is roughly two-thirds that for Grade 10, while for II's it is slightly over half that for Grade 10. The percentage of I's scoring at the mean decreases consistently from grade to grade; for II's it increases markedly from Grade 10 to 11, then slightly from Grade 11 to 12. This movement from below average to average achievement and later movement from average to above average achievement represents significant progress for students in average-below average groups.

While both Tables 10 and 11 reflect unequal degrees of progress between Grades 10 and 11 and Grades 11 and 12, with some categories even showing a slump at Grade 11, the figures do indicate improvement by students in both kinds of classes over the three-year span. The departure from the general trend noted at Grade 11 may--though we cannot be sure of this--be a reflection of a sampling error, since it cannot be assumed that all I or II sections at any grade level are exact equivalents.

DESPITE SOME IRREGULARITIES IN THE DEGREE OF YEAR-TO-YEAR PROGRESS, THE FACT THAT GROUP PROGRESS IS EVIDENCED BY THE SCORES OF TWELFTH GRADERS IN BOTH AVERAGE-ABOVE AVERAGE AND AVERAGE-BELOW AVERAGE CLASSES SUGGESTS THAT OUR APPROACH TO WRITTEN COMPOSITION IS SOUND. OUR CURRICULUM IS BASED ON THE CONVICTION THAT EXPOSITORY COMPOSITION INVOLVES A PROCESS WHICH CAN BE DESCRIBED AND TAUGHT AND WHICH REQUIRES THE APPLICATION OF VARIOUS SPECIFIC SKILLS THAT CAN BE MASTERED TO SOME DEGREE BY STUDENTS AT ALL ABILITY LEVELS.

If this conclusion is valid, it should be supported when internal, local achievement is judged against the external standard provided in the norms tables for the STEP Test. These national norms, according to the Educational Testing Service, are based on national samples "scrupulously chosen to be representative" and thus reflect the typical achievement of students at the various grade levels in other schools throughout the country.

It should be remembered that group scores on the STEP Writing Test may be assumed to reflect the influence of instruction students have received. Differences in native ability, motivation, home background and breadth of reading experience might, of course, be factors in individual writing progress. However, since the national norms represent typical achievement of groups carefully selected to be representative, since our large actual test samples represent all but a small minority of the lowest achievers at each grade level, and since the group scores of these actual samples were adjusted to compensate for this exclusion, the effect of these uncontrollable variables on group averages would be minimized. Maturation is not a factor; if higher scores at higher grade levels were due only to the natural process of growing up, we would expect less variation between the group scores at the same grade level in different schools.

The "influence of instruction" may not be strictly limited to the effect of direct instruction in written composition. Instruction in other language arts areas may also affect writing competence. The literature and language strands in our curriculum and the deliberate integration of learning experiences in these areas with the teaching of written composition which distinguishes our local curriculum from the fragmented treatment that characterizes many language arts programs may also be factors. But it should be recalled (Tables 5-8) that local group scores on the standardized test exhibit the same general growth trends as those on the local measure, a performance test structured to require application of the process and skills emphasized in the written composition strand of our own curriculum. Therefore, if local STEP scores compare favorably with those of students receiving instruction elsewhere, the nature of the learning experiences specified in this strand of our local curriculum must be recognized as appropriate to our goal of improvement in writing competence and effective in helping students of various ability levels to progress toward that goal.

COMPARISON OF LOCAL SCORES WITH NATIONAL NORMS

The figures in the next three tables include all those relevant to a comparison of local scores with national norms available in the publisher's tables. Achievement by Richfield students is represented by two sets of figures. One represents the actual scores achieved by students in classes included in the test population, a large sampling of the total student population following the regular curriculum. The other (adjusted) figures are hypothetical scores computed to compensate for the exclusion of students in the nongraded classes from the test population. This was done by adding to the list hypothetical scores for the same percentage (25.1% for Phase A, 28.2% for Phase B, 27.6% for Phase C) of these slow-learning students as the proportion of the total class population these nongraded classes represent. By assuming that all of these students, had they taken the test, would have received the lowest possible converted score of 247 (though a few might perhaps have done somewhat better than this) and adding these hypothetical scores to those of students actually writing the test, these less able students were given proportionate representation in the "adjusted" scores. In fact, this type of student was perhaps more than adequately represented in the adjusted sample; we know that some students who should be in these classes are actually in regular sections.

Since our students wrote the STEP Test in March and most of the norms tables are based on fall testing, we took an added precaution when selecting the particular sets of grade-level scores with which to compare the achievement of Richfield students in the three senior high school grades. In each instance, we compared local scores with national norms for fall testing at the next higher grade level. Where information was available, we also made comparisons with norms for spring testing and with the statistically projected, estimated norms for January-February testing of urban students. In interpreting the significance of the Grade 12 figures, it is important to remember that, while college students would take a different form of the test, the converted scores from test to test are comparable, according to the publishers. Nevertheless, such a comparison would tend to put our students at a disadvantage, since college students would constitute a more select group than high school seniors, as the publishers are careful to point out.

It is interesting also to note that the highest Richfield score was 342 at Grade 12, 338 at Grade 11, and 331 at Grade 10. The converted score for an errorless paper would be 350, but the highest achieved score listed for Grade 14 is in the 326-327 interval. In Richfield thirty students at Grade 12, seven at Grade 11 and five at Grade 10 scored at or above 326.

TABLE 12

Distribution of Local and Norms Group Scores on STEP Writing Test

Group Tested	Testing Time	Upper Quartile	Median (Midpoint)	Lower Quartile
Richfield Grade 10 (Actual)	March	309	297	286
Richfield Grade 10 (Adjusted)	March	309	297	283
Norms Group Grade 11	Fall	294	283	271
Urban Grade 11 (Estimated)	Fall	307	291	278
Richfield Grade 11 (Actual)	March	311	299	289
Richfield Grade 11 (Adjusted)	March	311	297	287
Norms Group Grade 12	Fall	299	287	277
Richfield Grade 12 (Actual)	March	317	306	295
Richfield Grade 12 (Adjusted)	March	317	304	293
Norms Group Grades 13-14	Fall	307	298	286
Norms Group Grade 12 Urban	Jan.-Feb.	310	295	283
Grade 12 (Estimated)	Spring	___[1]	290.5	___[1]

[1]A blank indicates that no such score appears in the available norms tables.

In every instance the converted scores that define the UQ, Median and LQ (the limits within which fall the top quarter, the upper half and the lowest quarter of scores) are higher for both Richfield's actual and adjusted sample populations at each grade level than for the norms group with which they are compared, though we might expect the opposite to be true in at least some cases.

TABLE 13
Percentages of Richfield Students Scoring
At or Above Various Medians and UQ Limits
and At or Below Various LQ Limits

Richfield Grade-Level Sample Group	Grade 11 Norms (Fall)			Grade 11 Urban Norms (Midyear)		
	UQ	Median	LQ	UQ	Median	LQ
10 (Actual Percentages)	58.6	72.9	7.05	28.6	67.4	15.4
10 (Adjusted Percentages)	58.6	73.84	10.97	27.4	64.6	18.9

	Grade 12 Norms (Fall)			Grade 12 Norms[1] (Estimated for Spring Testing)
	UQ	Median	LQ	Median
11 (Actual Percentages)	53.64	82.73	10.91	75.91
11 (Adjusted Percentages)	51.30	79.13	14.35	72.61

	Grade 11 Urban Norms[1] (Midyear)	Grade 12 Urban Norms[1] (Midyear)
	Median	Median
11 (Actual Percentages)	75.91	60.91
11 (Adjusted Percentages)	72.61	58.26

	Grades 13-14 Norms (Fall)		
	UQ	Median	LQ
12 (Actual Percentages)	49.34	70.48	9.25
12 (Adjusted Percentages)	47.26	67.51	13.08

	Grade 12 Norms[1] (Estimated for Spring Testing)	Grade 12 Urban Norms (Midyear)		
	Median	UQ	Median	LQ
12 (Actual Percentages)	85.46	43.17	77.09	7.49
12 (Adjusted Percentages)	81.86	41.35	73.84	11.39

[1]Quartile limits are not defined in the available tables.

In examining the preceding figures, it is important to remember (1) that our actual test sample was hypothetically adjusted to include a number of nongraded students proportionate to the percentage of the total class population represented by these students at each grade level, (2) that these students were all assumed to have earned the lowest score possible on the test and (3) that scores for each grade level have generally been compared with those of groups who might be expected to achieve at a higher level. In the first section of Table 13, for example, Richfield tenth graders tested in March are compared with the national norms for eleventh graders tested in fall; yet 58.6% of our students (more than double the expected 25%) scored in the upper quartile; 79.13% (more than one and one-half times the expected 50%) scored at or above the median; and 10.97% (less than half the expected 25%) scored within the lower quartile limits.

TABLE 14
Comparison of Richfield Mean Converted Scores
with Norms Group School Means

	Richfield Test Sample Mean (Actual-March)	Richfield Adjusted Sample Mean (Hypothetical-March)	Average School Means (Fall)	Highest Mean Achieved by Any School in the Means Group
Grade 10	297	295	277	287 (1 school)
Grade 11	299	297	281	291 (1 school)
Grade 12	306	304	285	295 (2 schools)

Here again it will be noted that the hypothetical Richfield means, which may possibly be lower than the actual means might have been if real scores for nongraded students had been available, are substantially higher than the norms school means for the next higher grade level. Our adjusted means for Grades 10 and 11 are, respectively, 14 and 12 points higher than the average school mean for fall testing at the next higher grade. Our adjusted means are 6-9 points above the highest means earned by students of the same grade in the one or two top ranking schools in the national sample tested in fall. It is clear from the norms tables that Richfield's adjusted mean for Grade 10 is in the 99-100 percentile range for Grade 11 of the norms schools. These tables indicate the highest average score earned by only two of the norms schools (295) as being at the 95th percentile. Since a one-point difference between means constitutes a jump of from one to thirteen percentile ranks elsewhere in these tables, Richfield's adjusted mean for Grade 11 (297) would be expected to place us <u>at the very least</u> in the 97+ percentile range for Grade 12.

No national means for Grade 13 are available, but Table 13 reflects a comparison of our twelfth graders with college freshmen and sophomores on the basis of median and quartile scores.

FROM ALL AVAILABLE EVIDENCE CONCERNING STEP WRITING TEST SCORES, IT APPEARS THAT THE ACHIEVEMENT OF RICHFIELD STUDENTS AT ALL GRADE LEVELS COMPARES VERY FAVORABLY WITH NATIONAL ACHIEVEMENT REFLECTED IN THE NORMS TABLES. IT HAS ALREADY BEEN POINTED OUT THAT SCORES ON THE LOCAL TEST OF ACTUAL PERFORMANCE ALSO REFLECT SUBSTANTIAL IMPROVEMENT FROM THE SOPHOMORE TO THE SENIOR YEAR. IT IS GRATIFYING THAT THE SCORES ON THE NATIONALLY STANDARDIZED TEST CORROBORATE THIS EVIDENCE OF GROWTH IN WRITING COMPETENCE DURING THE SENIOR HIGH SCHOOL YEARS. THIS REINFORCES THE VIEW THAT OUR WRITTEN COMPOSITION CURRICULUM, WITH ITS EMPHASIS ON TEACHING THE PROCESS AND SKILLS OF EXPOSITORY WRITING, IS SOUND IN CONCEPTION AND WORKABLE IN APPLICATION WITH CLASSES OF VARYING ABILITY.

ANALYSIS OF TEACHERS' RATINGS OF TEST COMPOSITIONS

Much has been written about the unreliability of teachers' grading of student compositions. The recurrent anecdote about the student who received A from one teacher and F from another on the same theme is not entirely without foundation in fact. Though one may wonder how he happened to submit the same paper to meet the requirements of two different assignments, the fact remains that readers differ greatly in their sensitivity to particular qualities in writing and in the weight they attach to them in grading. Numerous studies indicate that this is true even of the select group of English teachers at both the secondary and college level who serve as theme readers for the College Entrance Examination Board. This fact, it will be remembered, was one reason for having each test composition rated twice and for using the sum of the two independent ratings as the final score for each paper.

Considering the chaos that might have been expected to result from the involvement of all our teachers in the rating process, the data in the following tables assume considerable significance.

TABLE 15
Incidence of Ratings Categorized by Degree of Consistency Between the Two Separate Evaluations

Sample Group	Total Compositions Written	Number Receiving Two Identical Ratings (HH, MM, LL)	Number Receiving Adjacent Ratings (HM or LM)	Number Rated At Both Extremes (HL)
Grade 10	222	138	78	6
Grade 11	219	126	84	9
Grade 12	218	128	80	10
Grades 10-12	659	392	242	25

TABLE 16
Percentages of All Compositions Receiving
Various Possible Combined Ratings

Percentage Receiving Two Identical Ratings		Percentage Receiving Adjacent Ratings		Percentage Rated at Both Extremes[1]	
HH	13.66				
MM	30.50	HM	20.03		
LL	15.33	LM	16.69	HL	3.79
	59.49		36.72		3.79

[1]These HL ratings were scattered among sets of papers rated by nine pairs of raters, or eighteen teachers; four pairs of raters, or eight teachers, produced none of them. These extreme ratings were excluded from the test results as unreliable measures of students' "real" achievement.

To express the relationship between the two sets of ratings on these test compositions in the form of a coefficient of correlation is deceptive. Even though it does indicate some degree of correlation rather than a lack of correlation or a negative correlation, the figure (.3028) nevertheless suggests a lower degree of reliability than it should because the statistical computation of such a correlation treats all differing ratings as representing disagreement to the same degree, making no distinction between extreme differences (H and L) and slight differences (H and M or L and M).

A truer picture results from analyzing the percentages of ratings representing the various degrees of agreement or disagreement. The number of cases where both ratings agree perfectly (59.49%) is over fifteen times as large as the number where they disagree entirely (3.79%). Though a substantial number (36.72%) were rated in two adjacent categories (H and M or L and M), their number is only slightly ov alf that for identical ratings (HM, MM, or LL) but almost ten times that for ratings at opposite extremes (H and L). Further, the degree of real disagreement represented by these adjacent ratings (HM and ML) is probably less than it appears to be.

It must be remembered that raters were obliged to observe arbitrary percentages in assigning H, M and L ratings. In order to finish with 25% at each extreme and 50% in the middle group, most raters would probably have had to readjust their original groupings by moving some borderline cases from one category to another. Decisions between these borderline cases would be difficult to make and sometimes hard to justify, for there would be no marked difference in their quality. Nevertheless, some papers would have to be moved. It is not surprising that one teacher might therefore assign some marginal papers (not as good as the obvious H's but better than the obvious M's) to the H group while another would

retain these particular papers in the M group and move others, also representing marginal achievement, to the H group. The same circumstance would also be likely to occur with the low borderline cases. Thus, the resulting HM and LM ratings reflect no serious disagreement.

That there is far less real disagreement than complete or partial agreement is evidence of greater consistency than one might normally expect under the circumstances. Considering the number of teachers involved and the differing backgrounds of education and experience represented among these twenty-four English teachers, it is encouraging if not actually surprising to find serious disagreement in fewer than four percent of the 659 cases. This speaks well for the evaluative standards of the department--standards which it seems not unreasonable to attribute at least partly to the influence of our local curriculum, which provides common goals and procedures for teaching--and consequently evaluating--written composition.

Other points of interest concerning the combined ratings which constituted the final scores on the test compositions are the resemblance the total distribution for grades 10-12 bears to the normal curve and the fact that the Grade 10 scores exhibit a slight skew toward the low and the Grade 12 scores a slight skew toward the high end of the achievement scale. Since the school mean score was 800 and the computed standard deviation 254, we would expect, if the distribution is normal, that 68% of the scores would fall between 546 and 1054 (1 S.D. above and 1 S.D. below the mean), with the remaining 32% divided equally between the upper and lower remainders of the curve.[1] For the total sample this is very nearly the case; 69.875% of the scores fall within the 600-1000 limits. (This is the closest approximation of the computed limits that is possible, since there were only five actually achievable combined scores.) Of the remaining scores, 14.195% fall above and 15.93% below this large middle group.

If we think of the total spread of all scores in the sample (exclusive of the 25 HL ratings discarded) in terms of letter grades from A to F, then this large middle group might be thought of as representing the range from B to D+, while the high group would include A to B+ and the low group D to F. Considering that the original ratings were limited to only three possibilities rather than to five letter grades with plus and minus combinations, the proportions come remarkably close to the normal distribution as described in ESSENTIALS OF MEASUREMENT FOR TEACHERS, by Walter Durost and George Prescott. They place 66% of the grades between the limits of what they term the considerably below average and the considerably above average segments of the curve, with 17% at each end. This distribution also tallies closely with the percentages designated in the class record book supplied

---

[1]Georgia Sachs Adams, MEASUREMENT AND EVALUATION (Holt, Rinehart and Winston, 1965), p. 22.

Richfield Senior High School teachers; the 17% at each end would come from the 7% A's and F's and a portion of the 24% of B's and D's indicated there as normal when a five-letter grading scale is used. The chances are that a third reading of the HL themes would have resulted in their placement in one of these extreme categories, thus bringing our distribution even closer to the normal curve.

As indicated earlier (Table 6), the percentage of scores falling between the 600-1000 limits are not greatly different for Grades 10 and 12 (68.98% and 67.79% respectively), but the extremes are nearly reversed (20.37% below and 10.65% above at Grade 10; 20.67% above and 11.54% below at Grade 12). This is to be expected when all papers are judged by the same standard, without regard to the fact that under the ordinary circumstances of classroom grading, a substantially higher level of achievement would be required of a twelfth grader than of a tenth grader in order to earn a grade of A.

IN GENERAL THE ANALYSIS OF TEACHERS' RATINGS PROVIDES ENCOURAGING EVIDENCE NOT ONLY OF STUDENTS' PROGRESS, BUT ALSO OF DEPARTMENT CONSISTENCY IN EVALUATING STUDENT COMPOSITIONS. IT SEEMS LIKELY THAT BOTH OF THESE ADVANTAGES DERIVE AT LEAST TO SOME DEGREE FROM THE COMMON GOALS AND PROCEDURES STEMMING FROM THE NATURE OF OUR ESTABLISHED CURRICULUM FOR WRITTEN COMPOSITION. [1]

ASSESSMENT OF THE EFFECTS OF THE USE OF LAY READERS TO ASSIST TEACHERS IN THE EVALUATION OF STUDENT COMPOSITIONS

The lay reader program went into operation during the latter part of the 1965-66 school year and was continued, still on a very limited basis, during the 1966-67 school year. In assessing its effects our first concern is again with the data gathered by means of the testing program already described. We begin with the null hypothesis that the use of lay readers to assist teachers in evaluating written compositions produces no significant effect on students' achievement in this vital area of the language arts curriculum. The information in the following tables will serve to test this hypothesis.

TABULATION AND INTERPRETATION OF RELEVANT TEST DATA

Scores on both the STEP Writing Test and the test composition have been tabulated separately for students in groups served by lay readers (LR's) in groups not so served (NLR's).

---

[1] Judgments of students, parents, counselors and administrators concerning this established curriculum appear in a later section of this report, INTERPRETATION OF RESPONSES TO QUESTIONNAIRES.

TABLE 17
Comparative Ranges and Distributions of Scores
On the Test Composition[1]
(after exclusion of HL ratings)

Sample Group	Sample Size	Percentage of Sample Group Earning Each Combined Rating				
		HH	HM	MM	LM	LL
10 LR	108	12.96	12.04	33.33	20.37	21.30
10 NLR	108	8.33	19.44	32.41	20.37	19.45
11 LR	106	10.38	24.53	33.96	16.04	15.09
11 NLR	104	12.50	20.19	31.73	19.23	16.35
12 LR	80	28.75	22.50	20.00	16.25	12.50
12 NLR	128	15.63	25.78	35.15	12.50	10.94
10-12 LR	294	16.32	19.39	29.93	17.69	16.66
10-12 NLR	340	12.35	22.06	33.24	17.06	15.29

[1]Corresponding percentages for total grade-level samples appear in Table 6.

TABLE 18
Comparative Ranges and Distributions of Scores
On the STEP Writing Test[1]

Sample Group	Sample Size	Highest Score	Upper Quartile	Median	Lower Quartile	Lowest Score
10 LR	114	331	309	296	284	254
10 NLR	113	328	311	299	287	250
11 LR	114	335	311	300	289	247
11 NLR	116	338	311	297	287	247
12 LR	83	342	322	311	299	262
12 NLR	144	338	315	304	295	268
10-12 LR	311	342	313	300	290	247
10-12 NLR	373	338	311	299	290	247

[1]Corresponding figures for total grade-level samples appear in Table 5.

TABLE 19
Comparative Achievement in Relation to the Mean
For the Entire School Sample on the Test Composition[1]

Sample Group	Group Mean	Percentage Scoring Above the School Mean (800)	Percentage Scoring At the School Mean (800)	Percentage Scoring Below the School Mean (800)
10 LR	750.00	25.00	33.33	41.67
10 NLR	753.7	27.77	32.41	39.81
11 LR	798.1	34.91	33.96	31.13
11 NLR	786.5	32.69	31.73	35.58
12 LR	877.5	51.25	20.00	28.75
12 NLR	845.3	41.41	35.15	33.44
10-12 LR	802.0	35.71	29.93	34.35
10-12 NLR	798.2	34.41	33.24	32.35

[1]Corresponding percentages for total grade-level samples appear in Table 8.

TABLE 20
Comparative Achievement in Relation to the Mean
For the Entire School Sample on the STEP Writing Test[1]

Sample Group	Group Mean	Percentage Scoring Above the School Mean (301)[2]	Percentage Scoring Below the School Mean (301)[2]
10 LR	302.47	35.09	64.91
10 NLR	297.5	38.94	61.06
11 LR	299.29	46.49	53.51
11 NLR	297.94	43.10	56.90
12 LR	309.5	66.27	33.73
12 NLR	291.34	47.92	52.08
10-12 LR	303.18	47.27	52.73
10-12 NLR	295.27	44.24	55.76

[1]Corresponding percentages for total grade-level samples appear in Table 7.

[2]No students scored exactly at the computed school mean on this measure.

If the use of lay readers exerts real influence for either good or ill on students' achievement in writing, then it seems reasonable to expect the scores at the various grade levels to reflect that influence in clearly differing tendencies of range and distribution for the LR and NLR groups. As Table 17 indicates, extreme scores on the test composition are identical for both groups at every grade level; the tabulation for each of the six groups includes both HH and LL ratings. Similar lack of a consistent distinguishing trend is apparent in the extreme scores on the STEP Writing Test (Table 18). The highest scores achieved by students in classes served by lay readers were three points higher at Grade 10, three points lower at Grade 11, and four points higher at Grade 12 than those earned by students in classes not provided with this service. On the other hand, while their lowest score was four points higher at Grade 10, it was six points lower at Grade 12 than that earned by students in non-lay reader classes and identical with theirs at Grade 11.

Analysis of the scores falling toward the opposite ends of the total distribution for test results on the two measures provides another basis for comparing extremes of achievement by the two kinds of groups. While the HH and LL ratings on the test composition do not represent the same percentage of students as do the scores falling within the upper and lower quartiles for the STEP Test, these two sets of data should give some indication of any consistent relationship of LR and NLR classes in terms of their highest and lowest groups of earned scores (Tables 17 and 18). The percentage of LR's earning the highest possible combined rating is slightly larger at Grade 10 and almost twice as large at Grade 12 as the percentage of NLR's, but at Grade 11 it is slightly smaller. However, the percentage of LR's earning the lowest possible rating is also slightly larger at Grades 10 and 12, and it is slightly smaller at Grade 11 than for the group at the same grade level. The score defining the point at or above which twenty-five percent of each group's STEP scores fall is two points lower for LR's at Grade 10, the same for both groups at Grade 11, and seven points higher at Grade 12. The score at or below which twenty-five percent of each group's scores fall is three points lower for LR's at Grade 10 but two points higher at Grade 11 and four points higher at Grade 12. Again no consistent distinction between lay reader and non-lay reader groups emerges.

Two measures of central tendency are also recorded in the preceding tables. The median or midpoint score on the test composition is identical for five of the six groups (800); the only exception is at Grade 12, where it would fall in the 1000 category, since over 50% of scores are above 800 (Table 19). The finer distinctions possible among STEP scores may perhaps partially account for the fact that on this test the median for LR's (Table 18) is three points lower at Grade 10 and three points higher at Grade 11 than that for NLR's. Although the median for LR's on this test as on the other is higher than that for NLR's at Grade 12 (7 points on the STEP Test), the relationship between LR's and NLR's is not consistent for all three grade levels.

The other measure of central tendency is more useful for our purposes. Since it is less affected by a few extreme scores than the median, the mean gives a better picture of the performance typical of a group. On both tests this average score (Tables 19 and 20) is higher for LR's than for NLR's

except at Grade 10, where the mean on the test composition is a little lower for LR's than for NLR's. At Grade 11 the test composition mean is considerably higher but the STEP mean is only very lightly higher for LR's; at Grade 12 the difference in points is considerable on both measures, but larger on the composition than on the STEP Writing Test. In terms of average scores, both sets of test results suggest superior achievement by students in classes served by lay readers at Grade 11 and 12 but slightly inferior achievement by these classes at Grade 10.

We can also judge the achievement of each group at each grade level by comparing it with the school means; that is, with the average score for all students in the total sample tested by each of the two instruments. The percentages of LR's scoring above the school means exhibit the same tendencies on both the STEP Test and test composition (Tables 19 and 20). It is slightly lower at Grade 10, slightly higher at Grade 11 and substantially higher at Grade 12 than that for NLR's.

The combined percentage of LR's scoring <u>above</u> and <u>at</u> the mean on the test composition is slightly lower at Grade 10 and somewhat higher at Grade 11, but somewhat lower again at Grade 12 than that for NLR's. The importance of the difference between the tendencies of the above-mean scores and the above-plus-at-mean scores is not clear. To the writer, however, since no scores fall exactly <u>at</u> the mean for the STEP Test, it would seem more logical to compare the above-plus-at-mean percentages for the test composition with the above-mean percentages for the standardized test. When this is the basis for comparison, the similarity of tendency noted in the previous paragraph disappears.

The same inconclusiveness characterizes the data for scores falling below the school means. The percentages scoring below the average for the total test sample is lower for LR's than NLR's on both testing instruments at Grades 11 and 12--though in differing degrees on the two tests and at the two grade levels--but slightly higher at Grade 10.

ANALYSIS OF THEIR SCORES AT EACH GRADE LEVEL YIELDS NO CLEAR EVIDENCE THAT EITHER THE LR OR THE NLR GROUP CONSISTENTLY OUTPERFORMED THE OTHER.

At this point it seems important to mention some variables in the operation of the lay reader program which may have influenced the scores of LR groups at the various grade levels. At the time of testing, with the lay reader program in its second year, fewer than one-fourth of the students in LR groups included in the test sample were having their second experience with it, while the rest were receiving this service for the first time. Even among this latter group, the actual experience would differ in quantity from class section to class section. For various reasons, not all teachers used lay reader service to the same degree or for the same portion of the school year.

The quality of the experience would also be likely to differ from class section to class section because of varying circumstances. Several readers were involved, some for the first time and some for the second. Of those in their second year with the program, some were working with the same teacher and some with a new partner. Some readers were quicker to complete

their part in the evaluation process than others. Some teachers were in their first year of involvement with the program; others in their second. The working relationship between reader and teacher was not equally satisfactory in all instances.

Sampling error may also have been a factor. Not only was considerable disparity in the size of LR and NLR test populations at Grade 12 unavoidable, but also there could be no certainty that at any grade level either the LR and NLR class sections tested represented in exactly the same proportion the various ability levels of students.

The effects of the operation of these variables on the test data for smaller groups may be expected to cancel each other to some degree when their scores are combined. For this reason it is perhaps more appropriate for our purpose to compare test results for the total sampling of students served by lay readers (LR's, Grades 10-12) with those for the total sampling of students not so served (NLR's, Grades 10-12). Information relevant to this sort of comparison appears in the last two lines of each of the four preceding tables.

The upper quartile STEP score for LR's (Table 18) and the highest earned score are slightly higher for LR's, while the lower quartile limit and lowest earned score are the same for both groups. The median on the test composition (Table 17) is the same for both groups; the STEP median for LR's is very slightly higher than that for NLR's (Table 18). Of all students scoring below the school mean on the test composition (Table 19) the percentage is slightly higher for those in LR than for those in NLR groups. The percentage scoring above this mean is also slightly higher for LR's, though the combined percentage of those scoring <u>above</u> and <u>at</u> the mean is slightly higher for the NLR's. On the STEP Writing Test, the percentage of students scoring above the school mean is 3.03% higher for LR's than for NLR's and, obviously, the percentage scoring below the mean is correspondingly lower for them than for NLR's (Table 20).

As has been pointed out earlier, the most revealing measure of the performance typical of a group is the arithmetic mean, or average of all scores earned by that group. Considering that the grade-level means on the test composition for LR and NLR groups exhibit a range of over 100 points (Table 19), the superiority of the total LR group reflected in the difference between the means of the two groups (3.8 points) is slight. The difference between the STEP means for the two groups (7.91 points higher for LR's) is not inconsiderable, on the other hand, since only 18.16 points separate the lowest from the highest mean for any of the six smaller grade-level groups.

COMPARISON BETWEEN THE TOTAL LAY READER AND TOTAL NON-LAY READER GROUPS, THOUGH IT REFLECTS SOME INCONSISTENCIES, PICTURES LESS ERRATIC SCORING TENDENCIES THAN THOSE OF THE SMALLER LR AND NLR GROUPS AT THE SEPARATE GRADE LEVELS. ON MOST POINTS OF COMPARISON PROVIDED BY THE DATA FROM BOTH SETS OF TEST RESULTS, THE ADVANTAGE IS ON THE SIDE OF THE STUDENTS SERVED BY LAY READERS. TRUE, IN ONE INSTANCE THEIR ACHIEVEMENT IS INFERIOR TO THAT OF THE GROUP NOT SO SERVED, AND IN SOME OTHERS IT IS MERELY EQUAL TO THEIRS, BUT IN MOST CASES IT IS SUPERIOR. HOWEVER, EXCEPT FOR THEIR SUBSTANTIALLY HIGHER STEP MEAN, THE DEGREE OF THIS

SUPERIORITY IS SLIGHT. FURTHER, IT IS IMPOSSIBLE TO TELL HOW MUCH THE SCORES MAY HAVE BEEN AFFECTED BY THE OTHER VARIABLES PREVIOUSLY MENTIONED.

FOR THESE REASONS OUR CONCLUSION IS THE CAUTIOUS ONE THAT NO SIGNIFICANT DIFFERENCE IN STUDENTS' WRITING ACHIEVEMENT RESULTS FROM OUR USE OF LAY READERS TO ASSIST TEACHERS IN THE EVALUATION OF WRITTEN COMPOSITIONS. WHETHER OR NOT THIS CONCLUSION IS OVER-CAUTIOUS, IT IS DEFINITELY NOT AN INDICTMENT OF THE LAY READER PROGRAM. WHILE WE CANNOT CLAIM THAT THE PROGRAM HAS PRODUCED STARTLING IMPROVEMENT IN STUDENTS' WRITING SKILLS AS MEASURED BY TWO TEST INSTRUMENTS, WE CAN AT THE VERY LEAST CONFIDENTLY ASSERT THAT IT HAS NOT ADVERSELY AFFECTED THEIR ACHIEVEMENT IN THIS VITAL AREA OF THE LANGUAGE ARTS CURRICULUM.

Our findings compare favorably with those reported in the available research literature relating to the use of lay readers. There have been two major studies in this field, one supported by SUPRAD (School and University Program in Research and Development) of the Cooperative Research Program of the U. S. Office of Health, Education and Welfare and the other set up by Educational Testing Service. Together these research projects appraised lay reader programs in nineteen communities located in sixteen states. Results of the SUPRAD study were reported in the November, 1961 ENGLISH JOURNAL. Both studies are summarized in THE LAY READER PROGRAM: BACKGROUNDS AND PROCEDURES, a monograph by Virginia M. Burke of the University of Wisconsin-Milwaukee from which the quotations in this paragraph have been taken. As evidence that "the project yielded positive results," the SUPRAD report states (p. 9) that "no critical losses in quality of student writing" occurred, that "its quality . . . followed a normal pattern of development, with students progressing at just about the speed that might have been expected" (p. 10). Designed "to determine whether lay readers could be used without significant loss of improvement in student writing" (p. 9) the results of the ETS project "tended to show that students with lay readers were not handicapped. While no extraordinary gains in writing were discovered, there seemed to be no serious losses" (p. 10). *THESE STATEMENTS FROM RESPECTED SOURCES LEND SUPPORT TO THE VIEW THAT OUR TEST RESULTS CONSTITUTE SUFFICIENT JUSTIFICATION FOR PROVIDING STUDENTS WITH LAY READER SERVICE.*

INTERPRETATION OF RESPONSES TO QUESTIONNAIRE ITEMS

Test results are not the only criterion for judging the effects of lay reader service even on students' writing achievement, not to mention the program's indirect consequences for other aspects of their learning experience, which will be dealt with later. A student's attitude toward writing and toward the kind of evaluation his compositions receive may have considerable long-range influence on his progress even though it may not produce immediately apparent improvement in measurable achievement.

Certain items on the Student Questionnaire (See appendix) were designed to reflect these attitudes. Analysis reveals a significant relationship between the responses to the items dealing with what students understood to be the purposes of the lay reader program and to those concerned with their attitudes toward continued participation in it.

Before proceeding to interpret the specific items, a word of explanation is in order. The number of responses to items in the questionnaire exceeds the number of students in the test sample. This is because all classes served by lay readers were asked to fill out questionnaires, but not all of them were included in the testing. There is also a difference in the total number of responses to various questionnaire items; not all students responded to every statement despite the request that they do so. Since most students participating in the lay reader program did, however, respond to all items and since the format of the questionnaire assured the anonymity of respondents--even to the point of preventing identification of any student by his handwriting--it may be assumed that a tally of the responses represents the true feelings of students toward their experience with the program. For the purposes of this report, the original five-point scale has in most instances been reduced to three: agreement with the statement (total of SA and A responses), disagreement with it (total of DA and D responses), and inability to decide between agreement and disagreement with it (U responses).

It is overwhelmingly clear from the tallies of these three varieties of responses that students believed it a major purpose of the program to help them improve their writing skills. On item 8, which states this purpose, 348 indicated agreement, with only 23 disagreeing and 48 undecided. Students also gave unmistakable evidence of their awareness that the emphasis of the lay reader program, as of our curriculum, is on improving writing quality rather than merely increasing the quantity of composition assignments; 270 disagreed with the opposite statement (Item 1), while 96 agreed and 26 were undecided. Responses to items 5 and 6 reflect students' recognition that thorough analysis and detailed comments are a part of the meaningful evaluation to which the lay reader program is intended to contribute. Of the combined responses to the two items, 696 register agreement with the statements as representing purposes of the program, while only 51 reflect disagreement and 23 indecision. The combined totals for the responses to items 23, 26 and 27 (783 in agreement, 293 in disagreement, 92 undecided) show that most students recognize the lay reader as contributing to this meaningful evaluation of their compositions. Students seem to be indicating that praise and encouragement, more conducive to improvement than exclusive attention to error or weakness, are part of effective evaluation and also that these qualities actually characterized the lay readers' comments. This is reflected in the attitude of students toward their own future efforts. Only 13 (Item 25) felt strongly that lay reader comments tended to be discouraging and 49 indicated some discouragement, while 283 were not discouraged (210 disagreeing and another 73 strongly disagreeing with the statement) and 48 were undecided. The reader's interest in students' progress is also a factor in evaluation that fosters improvement. Students' responses to item 24 (188 agreeing, 114 disagreeing, 84 undecided) suggest that even though limited to written communication, the lay readers have made their concern evident to more students than one might expect.

WHEN ONE CONSIDERS STUDENTS' APPARENT AWARENESS OF THE PURPOSES OF THE LAY READER PROGRAM AND OF SIGNIFICANT ASPECTS OF ITS OPERATION, IT SEEMS CLEAR THAT THE ATTITUDE OF MOST STUDENTS TOWARD IMPROVING THEIR

WRITING SKILLS IS A POSITIVE ONE. IT IS ALSO APPARENT THAT THEY RECOGNIZE THE ASSISTANCE THEY HAVE RECEIVED THROUGH A PROGRAM DESIGNED TO CONTRIBUTE TOWARD MORE EFFECTIVE EVALUATION OF THEIR COMPOSITIONS--AN IMPORTANT AID TO IMPROVEMENT IN THE FUTURE (SOMETIMES EVEN THE RATHER DISTANT FUTURE, AS HAS BEEN POINTED OUT ABOVE).

The difference between the number agreeing with the statement in item 7 (182) and those disagreeing with it (155) was smaller than expected, since teachers themselves consider this special provision practical evidence of official concern with fostering maximum growth in writing competence. Perhaps students believed that their recognition of the importance of writing well was already assumed by school officials and therefore not a purpose of the program. At any rate, the number who did see this as a motive for the expenditures necessary to provide lay reader service is substantial.

The number indicating agreement with statement 16 is also smaller than was expected (159 as against 163 expressing disagreement). From the reactions of participants in lay reader programs in a number of communities, Miss Burke concludes in her previously mentioned monograph that such a program "takes writing out of the 'exercise' category and makes it a serious business of communication to the world outside the classroom" (p. 25). Perhaps use of the word *author* in the statement made agreement seem presumptuous. Perhaps students thought of the lay reader as another teacher. Yet communication with an unknown outsider was overwhelmingly recognized as a purpose of the program; 348 agreed with statement 8, while only 23 disagreed and 14 were undecided. Communication with someone other than his teacher (who often can read into a composition from his familiarity with the writer more than the student actually committed to paper and who may sometimes unknowingly do so) makes the student rely on the written word alone to say exactly what he means. Responses to items 8 and 17 (where 399 indicated agreement with the statement and 100 disagreement) reinforce each other and also, perhaps, the already substantial number of affirmative responses to statement 16.

Personal contacts between students and lay readers, including individual conferences, are frequently mentioned as an advantage of lay reader programs as they operated in other communities. Our program, in contrast, deliberately preserves the anonymity of students and readers for reasons already mentioned and also to emphasize that the teacher has not abdicated responsibility for or interest in the progress of each student. In Richfield, the attempt is also made to provide adequately for the necessary conferences about compositions that have been evaluated and returned, but they are conferences between teacher and student rather than between lay reader and student. The even division between students expressing agreement and disagreement with statement 19 (162 to 161) suggests that there is need for still greater provision for such conferences. These sessions are too time-consuming to be handled adequately during the regular class period without sacrificing some other valuable activity--a particularly undesirable circumstance because of the amount of class time required for

use of the laboratory method during the actual writing of compositions. It is also impossible to arrange for any considerable number outside the class day or outside the regular class period, when teachers' time is committed to teaching or supervising students for four more periods and to planning future lessons during the remaining period. Conference opportunities will not be completely adequate until (1) allowance for the necessary time and designation of a suitable place for such meetings with students become a part of each teacher's schedule and (2) students' schedules are sufficiently flexible so that they can meet with their teachers during these specified times.

Students apparently recognized that the lay reader program is intended not to supplant but rather to augment their teachers' efforts in the evaluation of assigned compositions. In this, too, the Richfield program differs from lay reader programs as they have operated in some other communities. Here the teacher reads <u>all</u> papers, even those already evaluated by a lay reader, and the efforts of the two evaluators are intended to complement each other. Of the 387 responses to item 13, 322 reflected students' appreciation of the second point of view provided by the lay reader <u>in addition to</u> the teacher's comments. On items 14 and 15, 500 responses indicated that the lay readers seemed to be aware of the teacher's aims and emphases in the composition assignments and to judge the results in terms of these requirements, so that generally there was agreement between the comments of the two evaluators. Understanding and agreement were not always complete, according to some students (169 indicated this view and 103 were unable to decide). Obviously, not all partnerships were equally satisfactory.

Nevertheless, the addition of a second reader to aid in evaluating their compositions has not caused most students to feel either confused or threatened. Of those responding to item 14, 267 felt no confusion, while 80 experienced some confusion and 38 were, strangely, uncertain about their reaction. Responses to item 18 indicate that while some students (101) at first felt uncomfortable about having a stranger comment on their writing, far more (259) did not. It is not clear whether the initially disquieting effect persisted for those who experienced it nor whether the 36 indecisive responses may represent early discomfort that later disappeared.

In more instances than not, the quality of lay readers' comments has justified their involvement in the evaluation of written compositions. That their contribution has been well accepted by students is indicated in some questionnaire responses already mentioned, and this impression is reinforced by responses to other items. In the opinion of most students, their evaluative comments were fair (269 agreed and only 49 disagreed with statement 20), courteous and friendly (295 agreed and only 48 disagreed with statement 21) and clear (Only 54 found them hard to understand, while 296 had no such difficulty).

Most students believed that the benefits to them were greater because they and their lay readers remained unknown to each other. Of response totals on items 11 and 12, 515 indicated a preference for relying solely

on written communication with an unknown reader who did not know them, while only 163 reflected a desire to meet and confer with the lay reader and 94 were undecided. The situation in Richfield seems to be in strong contrast with what has occurred in some other communities where the identity of lay readers was not made known to students. In the publication referred to earlier, Miss Burke makes the following statement (p. 21):

> When students are exposed to a challenge at which even adults quail--submitting writing to an unknown editor--they become angry and anxious about 'the unknown godlike creature passing judgment' on their themes. Under this system, conferences, one of the greatest values of the program, are impossible. To introduce elements of mystery and threat into a segment of their education which students are coming to believe is vital is a destructive procedure . . . .

Whether such negative actions are forestalled here by the quality of lay reader comments, by the students' awareness of the close working relationship between readers and teachers, by the provision for student-teacher conferences already mentioned, or by a combination of these and perhaps still other factors, the procedure as it has worked out in Richfield cannot be deemed destructive.

The reactions of our students as reflected in Table 21 compare favorably with those of students involved in a lay reader program in Racine, Wisconsin, reported--also by Miss Burke--in THE LAY READER PROGRAM IN ACTION, SPECIAL BULLETIN NO. 1 of the Wisconsin Council of Teachers of English, published in 1960. Of the 148 students who completed questionnaires, 60 indicated as "positive effects of the program on their work . . . that they now watch sentence structure more carefully, do more dictionary work, and organize, proofread, and revise more carefully" (p. 5).[1]

---

[1]The results of a study conducted in Sheboygan, Wisconsin and reported in the same publication are similar. Of the 247 students surveyed, 63 stated that they exerted greater care in their writing because of having a lay reader (p. 10). It is only fair to mention that classes involved in both the Wisconsin studies were limited to superior college-bound students, for whom it was perhaps already customary to take considerable pains with their composition assignments.

TABLE 21
Effects of Having a Lay Reader on Students' Effort

Questionnaire Item	Aspect of the Composition Process	Responses Indicating Degree of Effect: Effort Greater SA	A	Effect Uncertain U	Effort Unchanged D	SD
28	Organization of ideas	33	150	38	141	23
35	Selection of supporting evidence	40	154	37	126	22
29	Clear statement of ideas	45	159	26	112	21
30	Accurate word choice	24	161	40	123	23
31	Use of dictionary to check spelling	16	82	39	216	30
32	Punctuation	22	125	40	184	20
33	Revision	22	142	37	158	22
34	Proofreading	23	145	36	154	22

On six of the eight items, it will be noted, the number of responses indicating strong agreement is at least slightly larger than that for strong disagreement. Of the responses indicating mild agreement or disagreement, the number is larger for agreement in four instances and for disagreement in the other four. Relatively few were undecided about any of these items. The exception to the general trend (Item 31) probably reflects one influence of the laboratory method used in our teaching of composition. When writing is done in the classroom, where dictionaries are available and students are encouraged to use them, most of them, it appears, do habitually attempt to avoid spelling errors.

The totals for agreement (1245) and disagreement (1151) with the other seven statements indicate that only a few more students exerted special effort than the number whose effort was no greater because another reader besides the teacher would participate in evaluating their work. Although the percentage of Richfield students who did take greater care in these matters is larger than that for students surveyed in the Wisconsin study (See figures on page 29 of this report), nevertheless, the fact that so many exerted no additional effort might be considered evidence that our lay reader program fails to provide the hoped-for stimulus. On the other hand, another interpretation seems at least equally reasonable. These negative responses may support an optimistic conclusion about the curriculum and method of instruction that are distinguishing features of our structured writing program.

Since the number who exerted no special effort for the lay reader is so large, it seems unlikely that the reason is apathy and unconcern on the part of all these students. Perhaps the reason is that for many

it has become habitual to exert their best effort on these aspects of composition because teachers are doing a good job of making students aware that such care is important and of guiding them during the actual writing process. This habit is less likely to be developed by any but the superior, highly motivated students in schools where there is no structured composition curriculum, where writing assignments are still haphazard rather than process-centered and oriented to sequential skill development, where students do their writing outside the classroom without benefit of guidance at the moment of need, and where evaluation consists mainly of assigning a grade to the completed composition with little of the sort of meaningful comment and positive suggestion that will encourage the student and help him to do better on his next writing effort.

Two questionnaire items serve as the ultimate test of student reaction to the lay reader program. While only 54 said they would prefer not to have a lay reader work with their next year's teacher and 34 were undecided, 289 said they would like their next year's class to have the services of a lay reader. Similarly, only 85 wanted their teacher alone to judge their compositions and 56 were undecided, while 272 preferred that their papers be evaluated by both a lay reader and their teacher. *HOW STRONGLY STUDENTS APPROVE BOTH THE PROGRAM'S PURPOSES AND THEIR EXPERIENCE WITH IT IS INDICATED BY THE COMBINED RESPONSES TO ITEMS 9 AND 10. DESPITE SOME UNEXPLAINED INCONSISTENCY BETWEEN THE TWO SETS OF FIGURES, THE TOTALS EXCLUSIVE OF INDECISIVE ANSWERS CONSTITUTE A BETTER THAN THREE-TO-ONE ENDORSEMENT. THIS POSITIVE ATTITUDE TOWARD A SERVICE WHOSE PURPOSES STUDENTS RECOGNIZE AND WHOSE OPERATION THEY UNDERSTAND SUGGESTS A PREDISPOSITION TOWARD FUTURE PERSONAL EFFORT TO WRITE BETTER THAT IS SUFFICIENTLY WIDESPREAD TO WARRANT CONTINUATION AND EVEN EXPANSION OF THE LAY READER PROGRAM. IF SO MANY STUDENTS HAVE FOUND THE EXPERIENCE VALUABLE, IT SEEMS LIKELY THAT OTHERS WOULD ALSO BENEFIT AND SHOULD THEREFORE BE PROVIDED WITH THE OPPORTUNITY. STUDIES IN OTHER COMMUNITIES HAVE LED TO THE CONCLUSION THAT LAY READER SERVICE IS OF INSUFFICIENT BENEFIT TO ANY BUT THE SUPERIOR COLLEGE-BOUND STUDENT TO WARRANT PROVIDING IT FOR OTHERS. BECAUSE THIS CONCLUSION DOES NOT TALLY PERFECTLY WITH OUR OWN FINDINGS AND BECAUSE WE DO NOT HAVE SPECIAL TRACKS FOR TERMINAL AND COLLEGE-BOUND STUDENTS, SUCH A LOCAL LIMITATION SEEMS NEITHER JUST IN TERMS OF PRESENTLY AVAILABLE EVIDENCE NOR POSSIBLE TO IMPLEMENT.*

To supplement the data gathered directly from students via the questionnaire, the counselors, the principal and the assistant principal were asked to supply whatever information they had concerning attitudes toward lay reader service and judgments about its value. (See appendix.) With a single exception, they reported that they had heard no comments of any kind from either students or parents. If no news is good news--and in this case it seems to be, since neither our students nor their parents are likely to be reticent about making their dissatisfaction known to counselors or administrators--then it seems clear that our lay reader program has been well accepted. The exception mentioned was the principal, who reported that the few students' comments he had heard could be summed up in the statement "It is worthwhile doing written work when you know it will be looked at and corrected in its entirety." He further

stated that"parents discussing the procedure felt the same way and also indicated that they thought their students were improving themselves in communication as a result."

Though it seems relevant to judgment of the lay reader program as an aid to thorough evaluation of students' compositions, the above statement was actually a response to the request for comment on our structured curriculum for written composition. Another administrator had this to say:

> I have had many comments in my contacts with parents in regard to our curriculum, and all have been favorable. Many of our former students drop by from time to time and discuss their college programs as well as their problems. They are all grateful that they were able to take part in the composition program while in high school.

A counselor reported that comments from 1966 graduates enrolled in Freshman Composition at the University were encouraging in comparison with those from earlier graduates. These students "felt more comfortable and were succeeding better and earning better grades" than students who had had little or no experience with the curriculum adopted in fall of 1964, and they didn't voice the usual pleas about "making the kids write, write, write." This is especially encouraging in that members of the class of 1966 themselves had had only two years' experience with what is a six-year cumulative sequence.

Another counselor, one whose own unusual competence in writing in addition to his ten years' work with college-bound seniors puts him in a better position than any of his colleagues to assess trends in students' performance, reported "a very noticeable difference in the application blanks filled out by seniors. The job is done better than it was before, particularly in instances where a paragraph is required . . . . There is no question about the fact that these read more smoothly, have fewer grammatical errors, hold together structurally much better, and are more succinct. It used to be a real trial to go through them all. Now it is more of a pleasure." He noted similar improvement in the College Entrance Examination Board writing samples required of applicants for admission to some colleges. As areas in which greatest improvement was apparent, he mentioned organization, appropriate usage, and both greater precision and wider variety in word choice.

Of those responding to the questionnaire, the remaining counselor limited his comment to the method employed locally in teaching composition:

> The supervised writing laboratory seems to be very helpful. The immediate exchange of ideas with the supervising teacher is invaluable to the student.

> I feel that this approach should be utilized even more fully. The only limiting factor seems to be the impossible task you have due to class size at this time.

It is interesting to note that the underlined comment should come from a counselor. Like teachers themselves, he is obviously aware that while it may be possible to confer with twenty-five students during a class period, it is not possible to get around a class of thirty or more.

The comment is also significant in that it points up a difference in both conception and operation between the lay reader program in Richfield and those in some other communities, where the chief aim has been to reduce the teachers' work load and where an increase in the number of writing assignments has often been an added purpose. In the attempt to offset excessive pupil loads, the lay reader has in some cases been given complete responsibility for evaluating all or some of the students' compositions. In Richfield, on the contrary, the teacher relinquishes none of his responsibility for students' progress. Those papers which are evaluated by the lay reader actually receive two careful readings--sometimes even a third to check suggested revisions. Though the teacher's part in the evaluation process may eventually require less time because of the previous work of the lay reader, this is not the case at the beginning of the partnership. At first, in fact, the demand on the teacher's time is even greater. He must confer with his reader to establish mutual understanding of assignments and criteria for evaluating the resulting compositions, often completely evaluating some papers from a set beforehand and evaluating others jointly with the reader for purposes of illustration. Because of the cumulative, sequential nature of the assignments and the reader's need to know what has been emphasized in the teaching of the literary work on which the assignment is based, this is essential even though all our lay readers have themselves taught English, several in our own school system.

Even when excellent rapport has been established between compatible partners who have worked together for a considerable period, whatever time is saved for the teacher is freed only for application to other tasks. For the intent of our local program is borne out in actuality; the gain goes not to the teacher but to his students.

PARTICIPATING TEACHERS' ASSESSMENT OF THE LAY READER PROGRAM

In the proposal for the first limited trial use of lay readers to assist teachers in the evaluation of students' compositions, the following were indicated as hoped-for outcomes:

1. Favorable student response to the valuable second point of view provided by the lay reader

2. More thorough evaluation, including the sort of meaningful specific comments that show students <u>how</u> to improve

3. Greater promptness in evaluating compositions and returning them to their writers

4. Improvement in the quality of learning experiences in written composition through the more careful planning that becomes imperative when papers are to be evaluated by someone besides the teacher himself

5. Increase in the number of opportunities for students to write compositions and to benefit from evaluative comments and suggestions for improvement

6. Facilitation of teachers' identification of needs for review or reteaching and for improvement in future planning

7. Increase in students' awareness of the importance of writing well and in the degree of effort they exert toward this end

8. Favorable student response to the challenge of communicating ideas successfully to an unknown reader solely through what he, the unknown student writer, has actually committed to paper.

Like most students, the teachers involved in the program are generally agreed that it has indeed produced the first two of the originally anticipated outcomes, though some experienced at least initial frustration and not all found their partners equally satisfactory.

The third outcome has not so consistently materialized. In one case the reader took even longer than the teacher would have to evaluate the papers. In most instances the papers were not returned much sooner than they would otherwise have been, but this is not too surprising when one considers that after receiving them from the lay reader the teacher still had to read through each one, perhaps altering or adding to the lay reader's comments, and assign it a final grade. It must be remembered, however, that when the papers were returned they had been evaluated thoroughly--and by two readers instead of one.

Teachers unanimously agree that their responsibility to make it possible for the lay reader to do his work well has led to more careful planning for students' learning experiences in written composition. In order to make clear to the lay reader the specific requirements of the assignment, the background from which it developed and the specific criteria for judging the resulting compositions, participants have found it wise to prepare written directions for each assignment, including procedures and guidelines and identification of the elements to be emphasized in evaluation, and to give copies to both students and lay reader. They have even found themselves writing the compositions according to these directions before assigning them to students, to make sure in advance that their plans are workable and to anticipate problems students are likely to encounter.

It will be recalled that most students did not believe an increase in the number of writing assignments to be a purpose of the lay reader program. Their reaction reflects the fact that thus far, at least, it has not been possible for teachers to assign more compositions and at the same time insure the more important outcome in terms of increased writing competence--the provision of better learning experiences of the carefully planned, carefully supervised, carefully evaluated sort more likely to produce improvement than large numbers of superficial assignments perfunctorily completed and perfunctorily graded.

The lay reader's tabulation of recurrent errors or weaknesses has in some instances simplified the teacher's task of ascertaining what remedial instruction is needed by an entire class or a segment of it. It has also sometimes served to reveal unforeseen flaws in the teacher's plan. Thus alerted, the teacher can improve the structuring of subsequent composition experiences.

It has already been mentioned that the questionnaire responses indicated no consciousness among a large number of students of any great change in their attitude toward the importance of writing or in the degree of effort they devoted to preparing their compositions, but that most students did seem to react favorably to the challenge in the last item in the preceding list. Teachers found it difficult to tell how much students' seriousness about improving their writing skills could be attributed specifically to the lay reader program and how much may have stemmed from other causes such as the emphasis the teacher placed on composition in implementing the curriculum or the motivational effects of successful experience and recognizable improvement toward which the laboratory method aims. Whether or not they believe that students were always favorably affected by the lay reader's contribution to the evaluation of their compositions, the participating teachers consider the objectivity and second point of view provided by the lay reader generally helpful as sources of additional perspective about students' performance.

One significant outcome of the lay reader program was not included among the anticipated advantages listed in the original proposal. Teachers find that once the partnership is working well, they are no longer forced to divide their time and effort between what has already been and what now is. When they can give full attention to the current work instead of splitting their energies between their teaching of one unit in class and their evaluation of compositions that culminated the previous one during their prepration period and after school, teachers can do a better job of planning and conducting meaningful learning experiences. *LAY READER SERVICE THUS CONTRIBUTES TO THE IMPROVEMENT OF INSTRUCTION NOT ONLY IN WRITTEN COMPOSITION BUT ALSO IN OTHER AREAS OF OUR CURRICULUM.*

For teachers as for students, the most pragmatic evidence of belief in the value of the lay reader program is their desire to continue participation in it. Since the very small beginning of the program late in the 1965-1966 school year, when only two teachers were provided with lay reader service and only a few more indicated their willingness to try it, no teacher has asked to withdraw from the program. On the contrary, some who were at first reluctant have since become participants. Eleven teachers volunteered

to experiment with the program in 1966-67, and in both succeeding years eight made use of at least a portion of a lay reader's time. The number of lay readers has meanwhile increased from two to five, with the only one to leave the program doing so because increasing family responsibilities made prompt return of papers impossible.

IT IS APPARENT FROM THIS RECORD THAT WHILE THERE HAVE BEEN SOME FRUSTRATIONS, THEY HAVE BEEN NEITHER FREQUENT NOR SERIOUS ENOUGH TO PREVENT OVERWHELMING ENDORSEMENT OF THE PROGRAM BY ALL THREE GROUPS INVOLVED: STUDENTS, LAY READERS AND--PERHAPS MOST SIGNIFICANT, CONSIDERING THEIR CONCERN FOR THEIR STUDENTS' PROGRESS--THE TEACHERS WHO HAVE PARTICIPATED IN ITS EARLY STAGES. BECAUSE THE COST PER PUPIL IS SO SLIGHT IN TERMS OF VALUE RECEIVED, THE EXPENDITURE NEEDED TO MAKE LAY READER SERVICE PERMANENTLY AND INCREASINGLY AVAILABLE SEEMS A PRUDENT INVESTMENT FOR A COMMUNITY CONCERNED ABOUT QUALITY EDUCATION.

SUMMARY OF CONCLUSIONS BASED ON THIS STUDY

Our first evaluative study, though admittedly not the most sophisticated in design or in treatment of the data, has yielded encouraging results:

1. Scores on both the STEP Writing Test, a standardized objective test designed to measure writing skills, and the impromptu composition, a performance test requiring application of the process and skills emphasized in our local curriculum for written composition, indicate that our students do indeed improve these skills substantially from grade level to grade level.

2. This improvement is general for all students following the regular curriculum. While students in average-above average classes consistently earn a larger proportion of high scores than those in average-below average classes, students in average-below average classes at Grade 12 earn fewer low scores and more average and high scores than do students in these sections at Grade 10.

3. At every grade level the mean and median scores achieved by Richfield students on the STEP Writing Test are well above the national norms which reflect the typical performance of students at the same or next higher educational level in representative schools throughout the nation.

4. This favorable comparison with an objective, external standard that reflects achievement related to educational instruction speaks well for the learning experiences provided students in Richfield. The nationally standardized measure is designed to test the skills of written composition, and the impromptu expository composition requires the application of the process and skills emphasized in our local curriculum.

Since data from the two measures exhibit corresponding tendencies, we may conclude that our structured writing program does indeed promote the attainment of the objectives for which it was designed.

5. The large size and representative nature of the test sample, the high percentage of teachers with classes in the sample, and the equivalent term of experience with the curriculum for all grade levels involved in the testing--these and other factors support the view that the test scores reflect achievement as influenced by the nature of our curriculum rather than by other possibly causative circumstances.

6. The two independent ratings of each test composition reflect a far greater degree of agreement between the evaluators than might have been expected. Since the raters included all twenty-four English teachers--a large group of individuals differing considerably in training and experience--this consistency may be viewed as another indication of the influence of our curriculum in setting common goals for teaching and common standards for evaluating written composition.

7. Achievement scores on neither the impromptu test composition nor the standardized objective test of writing skills reflect any disadvantage to students from the use of lay readers to assist teachers in the evaluation of compositions--though there is no dramatic evidence, either, that the service is responsible for students' growth in writing competence. There is, however, clear indication from student and teacher participants, from lay readers, and from counselors and administrators that the program has been well received and that its value has been recognized. This value extends beyond the improvement of learning experience directly related to written composition to the improvement of learning activities related to other aspects of our language arts curriculum.

8. The department effort involved in this study was worthwhile. The results provide teachers some gratification for past achievement and encouragement for efforts to improve upon it in the future. Perhaps as important as the product was the experience of participating in the process; concerted effort of this sort is valuable in promoting dialog, self-evaluation and the broadened perspective that comes from involvement with others in a coopeative enterprise.

RECOMMENDATIONS STEMMING FROM THE FINDINGS OF THE STUDY

Continuing evaluation of the effectivensss of curriculum and instruction is a requisite for a successful educational program. This long-neglected responsibility of educators is receiving strong emphasis throughout our nation since education has become its biggest "business" establishment. Local stress on evaluation stems in part, certainly, from this national trend. But within the language arts department, whose members have contributed

so much to the creation and implementation of our local sequential curriculum, there is a special concern and a sincere desire to employ evaluation as a useful instrument in their constant effort to improve students' learning experiences in this crucial subject area.

Recommendations for suitable enabling action follow:

1. A study to assess students' growth in writing competence as an indication of the effectiveness of our structured curriculum for written composition and of the laboratory method emphasized in its teaching should be conducted at the senior high school level in the spring of 1968.

2. Similar studies should be carried out, if not annually, at least at frequent intervals thereafter.

3. In 1968-69 or as soon afterward as possible, a similar program of evaluation should be implemented at the junior high school level.

4. As soon as the necessary assistance is available, planning should begin for a follow-up study on the college performance of our recent graduates to determine how effectively our composition curriculum and instructional methods prepare Richfield students to compete successfully with graduates of other secondary schools. The study itself should be conducted as soon as possible, but at least by the 1970-71 school year, when members of the first graduating class to have completed the entire six-year sequence of the Richfield curriculum first implemented in 1964-65 will be college freshmen.

An ongoing program of evaluation would be advantageous in a number of ways, such as the following:

a. Additional studies would provide a sounder basis for judging the effectiveness of our written composition curriculum and its instructional implementation. If later results confirm our present judgments, we will have substantiation for continued confidence in the philosophy and methodology of our present writing program and encouragement to make it increasingly effective as we learn from longer experience with it. If the results of later studies conflict with our initial findings and cast doubt upon our present cautiously optimistic conclusion, we will be spurred to seek the underlying causes. We will need to reexamine not only our testing program and our individual teaching efforts but also our established composition curriculum and our approach to the teaching of writing skills. Results from the recommended follow-up study may also require us to consider whether alterations in policy concerning such matters as class size, scheduling, and teacher assignment may further our attempts to meet the special needs of the college-bound student, if they are indeed special--as they

may well be in degree if not in kind.

b. Alteration of the methods used in the first study for selecting the sample populations and tabulating the data may enable us in the next study to assess the progress and the relative standing of students in the nongraded sections and, possibly, to form some tentative judgment about the effectiveness of the experimental adjustment in grouping initiated in 1967-68 at the senior high school level.

c. Development of local norms for the STEP Writing Test would make it possible to compare (1) the achievement of future classes with that of classes constituting the norms groups and (2) the achievement of individual students with what is typical for students participating in the same kinds of learning experiences. This comparison of individual scores with local norms might be useful for scheduling purposes, particularly if we move to some more highly individualized form of programming.

d. Participation in this sort of evaluative study sharply focusses the attention of all department members on the common goals and procedures of our composition program and on the application of appropriate standards for evaluating all student compositions. This is vital to effective, consistent implementation of the official curriculum in a department with constantly changing membership.

e. Authorization of the modest expenditures required to conduct further studies would not only insure the greatest possible use of the work already done in the first one; it would give teachers positive evidence of administrative concern and support for a project department members consider sufficiently important to the best interests of their students to warrant their personal expenditure of additional time and effort beyond what is required of them.

The four recommendations thus far made have concerned assessment of students' achievement in written composition as related to the nature of our curriculum and the laboratory method of teaching. The final two concern the lay reader program:

5. Lay reader service should be continued and expanded to benefit as many students as possible.

6. Since a further attempt to evaluate the service on the basis of test data seems unlikely to prove any more conclusive than that reported here and since the beneficial effects of the lay reader program transcend objective measurement, no reevaluation of this aspect of our departmental operation should be undertaken, at least for the present.

IMPROMPTU TEST COMPOSITION

General Directions (Read silently as your teacher reads aloud. You will receive no other instruction. Your teacher will add no explanation or comment.)

1. Write your composition in ink on the unlined paper provided.

2. Instead of the usual heading, write only your 6-digit identification number in the upper righthand corner.

3. Use only one side of the paper. Additional sheets are available in case you need more than one page for your composition.

4. You may use this direction sheet for scratch paper. It is to be turned in at the end of the period.

5. Do not plan to write a rough draft. You will not have time to recopy it.

6. Dictionaries are available for your use.

7. The reader who judges your composition will be more concerned with ideas and organization than with spelling and punctuation and will consider quality more important than quantity.

Specific Directions (Read carefully. Your teacher will not read aloud.)

The material below will provide the subject for your composition and tell you what you are expected to do with this subject.

"Everybody wears a mask."

The "masks" people wear to hide their true nature or feelings affect their relationships with other people. In what ways and to what extent do people benefit or suffer because they and others conceal the truth about themselves?

. . . . . .

Express your ideas in a brief, well-planned composition. Support your views with examples drawn, if possible, from your own experience or observation rather than from your reading. (See IMPORTANT EXCEPTIONS below.) After you have finished your composition, underline the sentence you think best summarizes your central idea.

IMPORTANT EXCEPTIONS

1. If you use any illustrations taken from your reading, they must NOT come from literature you have studied in Senior High.

2. Do NOT mention your age or anything else that would reveal what grade you are in or whether the compositions you have written this year have been evaluated by both a lay reader and your teacher.

Circle: Grade 10 11 12

STUDENT QUESTIONNAIRE: REACTION TO THE LAY READER PROGRAM

During this school year some of your written compositions have been evaluated by a lay reader. This questionnaire has been designed to find out what you think and how you feel about this experience. The survey has a serious purpose, and its success depends upon your serious, thoughtful responses to the numbered statements below.

This is not a test. There are no "right" answers--only answers that are right for you because they represent your honest feelings and judgments. All responses are to remain anonymous. Not even your teacher is to know which paper is yours; therefore you are to omit the usual heading and simply circle the proper number in the upper righthand corner of this sheet to indicate what grade you are in.

Please read each statement carefully and then draw a circle around the symbol that best represents what is true for you. (Note that the statements are numbered at the right rather than at the left.)

PLEASE RESPOND TO EVERY ITEM.

In the first three items, circle either YES or NO.

My last year's English class had the services of a lay reader. 1. YES NO

The reasons for having a lay reader program are not clear to me. 2. YES NO

My teacher explained the role of the lay reader. 3. YES NO

In the rest of the questionnaire, respond as follows:

Circle SA if you strongly agree with the statement.
Circle A if you agree but not strongly.
Circle U if you are undecided or neutral. (There should be very few cases where you can't make up your mind.)
Circle D if you disagree but not strongly.
Circle SD if you disagree strongly with the statement.

I believe that a purpose of the lay reader program is to make it possible for students to have more opportunities to write compositions. 4. SA A U D SD

I believe that a purpose of the lay reader program is to make possible a more thorough analysis of students' compositions. 5. SA A U D SD

I believe that a purpose of this program is to make it possible for students to receive detailed comments on their compositions instead of only a grade and perhaps a brief general comment. 6. SA A U D SD

I believe that a reason for providing lay readers is to show us that not only English teachers but school administrators and board members believe that learning to write well is very important to us.	7.	SA	A	U	D	SD
I believe that a reason for having lay readers is to help us improve our writing skills by making us depend entirely on what we actually put on paper in order to communicate with someone we don't know and who doesn't know us.	8.	SA	A	U	D	SD
I would like to have a lay reader work with my teacher again next year. (Seniors, please answer as if you were going to be in school next year.)	9.	SA	A	U	D	SD
I would rather have only my teacher read and judge my papers.	10.	SA	A	U	D	SD
If a lay reader does read my papers, I would benefit more from the service if I could meet and confer with him.	11.	SA	A	U	D	SD
I would rather not meet my lay reader; he can judge my papers more objectively if we do not know each other.	12.	SA	A	U	D	SD
I like having a lay reader because I appreciate getting a second point of view in addition to my teacher's comments.	13.	SA	A	U	D	SD
I became confused by having two people make comments on my paper because they didn't always seem to agree.	14.	SA	A	U	D	SD
The lay reader seemed to know exactly what my teacher expected me to do in each writing assignment and judged my papers accordingly.	15.	SA	A	U	D	SD
Having a lay reader has made it seem a little as if I were an author writing for a "real" reader instead of just doing school work.	16.	SA	A	U	D	SD
Having a lay reader has made me realize that in order to communicate on paper I must take special pains to say exactly what I mean.	17.	SA	A	U	D	SD
At first I felt uncomfortable about having a stranger comment on my writing.	18.	SA	A	U	D	SD
I feel I was given sufficient opportunity to discuss the lay reader's comments with my teacher.	19.	SA	A	U	D	SD
The comments of my lay reader have been fair.	20.	SA	A	U	D	SD
The comments of my lay reader have been courteous and friendly.	21.	SA	A	U	D	SD

I have found it hard to understand the comments the lay reader wrote on my papers.	22.	SA	A	U	D	SD
The lay reader has usually made favorable comments as well as unfavorable comments about my written compositions.	23.	SA	A	U	D	SD
Although I don't know the lay reader, I feel he is interested in my writing progress.	24.	SA	A	U	D	SD
The comments of the lay reader have tended to discourage me.	25.	SA	A	U	D	SD
The lay reader has usually given reasons for his comments on my work.	26.	SA	A	U	D	SD
The lay reader has usually written a helpful summary comment about my entire composition besides marking errors and commenting on certain parts of the paper.	27.	SA	A	U	D	SD
Because I knew my papers might be read by a stranger, I made a greater effort to organize my ideas well.	28.	SA	A	U	D	SD
Because I knew my papers might be read by a stranger, I made a greater effort to state my thoughts clearly.	29.	SA	A	U	D	SD
Because I knew my papers might be read by a stranger, I have chosen my words more carefully.	30.	SA	A	U	D	SD
Because I knew my papers might be read by a stranger, I have checked my spelling with a dictionary.	31.	SA	A	U	D	SD
Because I knew my papers might be read by a stranger, I have been more careful in punctuating my sentences.	32.	SA	A	U	D	SD
Because I knew my papers might be read by a stranger, I have revised my work more thoroughly.	33.	SA	A	U	D	SD
Because I knew my papers might be read by a stranger, I have proofread them more carefully.	34.	SA	A	U	D	SD
Because I knew my papers might be read by a stranger, I have been more careful to select evidence that would really support my statements.	35.	SA	A	U	D	SD

ADDITIONAL COMMENTS

If you have an opinion or suggestion about the lay reader program which is not stated in any of the above items, write it on the back of this sheet.

A REQUEST FROM THE SENIOR HIGH SCHOOL LANGUAGE ARTS DEPARTMENT

TO:

We are attempting to evaluate two of our department enterprises, and we need your cooperation. We believe that as a counselor or principal you may possess evidence that would be a significant addition to the data we have collected by other means. Will you please help us by supplying the information requested on these sheets and returning them to any one of us as soon as you conveniently can?

Thank you.

Harold S. Webster, Evaluation Committee Chairman

Lucille Duggan, Department Chairman

Lillian B. Ford, District Chairman

<u>The Lay Reader Program</u>

This year and for a short time last year, lay readers have been working with some language arts teachers in the evaluation of written compositions. You may have heard comments from students or parents concerning this program. In the space below or on the back of this sheet, please record any evidence--either favorable or unfavorable--you may have concerning attitudes toward this innovation and judgments about the value of this experience to the students involved.

<u>The Written Composition Program, Grades 7-12</u>

Our present seniors are the first group to complete the senior high school portion of the 7-12 composition program initiated with the adoption of our LANGUAGE ARTS CURRICULUM GUIDE in September, 1964. You may have heard comments about this program from students, alumni, or parents; you may also have information from your own observation or from outside sources about the achievement of Richfield students in this vital area. In the space below or on the back of this sheet, please record your judgment concerning the effectiveness of the program thus far and any clues that might help us identify its particular strengths or weaknesses as a means to improvement of written composition skills.